(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 January 2001 (18.01.2001)

PCT

(10) International Publication Number
WO 01/04284 A1

(51) International Patent Classification[7]: C12N 15/00

(21) International Application Number: PCT/US00/19007

(22) International Filing Date: 13 July 2000 (13.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/143,711 13 July 1999 (13.07.1999) US

(71) Applicant *(for all designated States except US)*: **MIDWEST RESEARCH INSTITUTE** [US/US]; 425 Volker Boulevard, Kansas City, MO 64110 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ADNEY, William, S.** [US/US]; 13190 West 21st Avenue, Golden, CO 80401 (US). **DECKER, Stephen, R.** [US/US]; 820 Greenwood Drive, Berthoud, CO 80513 (US). **LANTZ McCARTER, Suzanne** [US/US]; 3072 West 39th Avenue, Denver, CO 80211 (US). **BAKER, John, O.** [US/US]; 18790 West 60th Avenue, Golden, CO 80403 (US). **NIEVES, Rafael** [US/US]; 1794 South Endicott Street, Lakewood, CO 80401 (US). **HIMMEL, Michael, E.** [US/US]; 9202 West Hialeah Place, Littleton, CO 80123-2148 (US). **VINZANT, Todd, B.** [US/US]; 16601 W. 15th Avenue, Golden, CO 80401 (US).

(74) Agent: **WHITE, Paul, J.**; National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, CO 80401 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published:**
— *With international search report.*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 01/04284 A1

(54) Title: CELLOBIOHYDROLASE REDUCED GLYCOSYLATION VARIANTS: CBHIN45A; CBHIN270A; AND CBHIN384A

(57) Abstract: The invention provides a method for making an active exoglucanase in a eukaryotic heterologous host, the method comprising reducing glycosylation of the exoglucanase, wherein reducing comprises replacing an N-glycosylation site amino acid residue with non-glycosyl accepting amino acid residue. The invention further provides a cellobiohydrolase, comprising the reduced glycosylation variant cellobiose enzymes CBHIN45A; CBHIN270A; or CBHIN384A, or any combination thereof.

# CELLOBIOHYDROLASE REDUCED GLYCOSYLATION VARIANTS: CBHIN45A; CBHIN270A; AND CBHIN384A.

## Technical Field.

This invention relates to exoglucanases. More specifically, it relates to *Trichoderma reesei* cellobiohydrolase I reduced glycosylation variants which enable expression of the active enzyme in a heterologous host.

## Background Art.

The surface chemistry of acid pretreated-biomass, used in ethanol production, is different from that found in plant tissues, naturally digested by fungal cellulase enzymes, in two important ways: (1) pretreatment heats the substrate past the phase-transition temperature of lignin; and (2) pretreated biomass contains less acetylated hemicellulose. Thus, it is believed, that the cellulose fibers of pretreated-biomass are coated with displaced and modified lignin. This alteration results in a non-specific binding of the protein with the biomass, which impedes enzymatic activity. Moreover, where the pretreated biomass is a hardwood-pulp it contains a weak net-negatively charged surface, which is not observed in native wood. Therefore, for the efficient production of ethanol from pretreated biomass it is desirable to enhance the catalytic activity of glycosyl hydrolases on acid hydrolyzed hardwoods.

*Trichoderma reesei* CBH I is a mesophilic cellulase enzyme, and comprises a major catalyst in the overall hydrolysis of cellulose. An artificial ternary cellulase system consisting of a 90:10:2 mixture of *T. reesei* CBH I, *A. cellulolyticus* EI, and *A. niger* β-D-glucosidase is capable of releasing as much reducing sugar from pretreated yellow poplar as the native *T. reesei* system after 120 h. This result is encouraging for the ultimate success of engineered cellulase systems, because this artificial enzyme system was tested at 50°C, a temperature far below that considered optimal for EI, in order to spare the more heat labile enzymes CBH I and β-D-glucosidase. In order to increase the efficiency of such artificial enzyme systems it is desirable to engineer new *T. reesei* CBH I variant enzymes capable of active expression in a heterologous host. The heterologous host *Aspergillus awamori*, could provide an excellent capacity for synthesis and secretion of *T. reesei* CBH I because of its ability to correctly fold and post-translationally modify proteins of eukaryotic origin. Moreover, *A. awamori* is believed to be an excellent test-bed for *Trichoderma* coding sequences and resolves some of the problems associated with direct site directed mutagenesis in *Trichoderma*.

In consideration of the foregoing, it is therefore desirable to provide variant cellulase

enzymes having enzymatic activity when expressed in an heterologous host.

Disclosure of Invention.

It is a general object of the present invention to provide variant cellulase enzymes having enzymatic activity when expressed in a heterologous host, such as a filamentous fungi or yeast.

Another object of the invention is to provide a variant exoglucanase characterized by a reduction in glycosylation when expressed in a heterologous host.

Another object of the invention is to provide an active cellobiohydrolase enzyme capable of expression in heterologous fungi or yeast.

It is yet another object of the invention to provide a method for reducing the glycosylation of a cellobiohydrolase enzyme for expression in a heterologous host.

The foregoing specific objects and advantages of the invention are illustrative of those which can be achieved by the present invention and are not intended to be exhaustive or limiting of the possible advantages which can be realized. Thus, those and other objects and advantages of the invention will be apparent from the description herein or can be learned from practicing the invention, both as embodied herein or as modified in view of any variations which may be apparent to those skilled in the art.

Briefly, the invention provides a method for making an active exoglucanase in a heterologous host; the method comprising reducing glycosylation of the exoglucanase, reducing glycosylation further comprising replacing an N-glycosylation site amino acid residue with a non-glycosyl accepting amino acid residue. The invention further provides a cellobiohydrolase, comprising the reduced glycosylation variant cellobiose enzymes CBHIN45A; CBHIN270A; or CBHIN384A, or any combination thereof.

Best Mode for Carrying out the Invention.

Unless specifically defined otherwise, all technical or scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are now described.

A method for reducing the glycosylation of an expressed *Trichoderma reesei* CBHI protein by site-directed mutagenesis ("SDM") is disclosed. The method includes replacing an N-glycosylation site amino acid residue, such as asparagines 45, 270, and/or 384 of SEQ. ID NO: 4

(referenced herein as CBHIN45A, CBHIN270A and CBHIN384A, respectively), with a non-glycosyl accepting amino acid residue, such as is alanine. Various mutagenesis kits for SDM are available to those skilled in the art and the methods for SDM are well known. The description below discloses a procedure for making and using CBHI variants: CBHIN45A; CBHIN270A; and CBHIN384A. The examples below demonstrate the expression of active CBH I in the heterologous fungus *Aspergillus awamori*.

Industrial Applicability.

Site-Directed Mutagenesis of *Trichoderma reesei* CBH I for Reduced Glycosylation.

*Aspergillus awamori* was transformed with various versions of the *cbhl* gene from *Trichoderma reesei*. *The cbhl* genes included both cDNA and genomic (intron containing) versions. These were altered by site-directed mutagenesis for the specific purpose of reducing the glycosylation of the expressed CBH I protein through replacement of the N-glycosylation site amino acid residues (asparagine) with non-glycosyl accepting amino acid residues (alanine). The gene was propagated in an *E. coli* vector plasmid (pPFE2) under the control of the *Aspergillus awamori* glucoamylase promoter and signal sequence, and *trpC* terminator, and carrying resistance to ampicillin *(E. coli* selection) and Zeocin (Bleomycin) *Aspergillus* selection. One altered rCBH I variant, CBHIN270A, SEQ. ID. NO: 2, was isolated from cultures and determined to be consistent with native CBH I, SEQ. ID. NO: 4, with respect to kinetics on *p*NPL and was only slightly higher in molecular weight. Thus, construction of the triple reduced glycosylation mutant CBH 1, CBHIN270A (SEQ. ID. NO: 2) / CBHIN45A (SEQ. ID. NO: 1) / CBHIN384A (SEQ. ID. NO: 3), may provide a viable means of producing active CBH I in heterologous fungal or yeast which do not require the cellobiose/lactose induction cascade, known in *Trichoderma.* It is believed that reduced glycosylation CBH I mutants would also serve effectively in yeast-based high throughput screens, which are normally rendered unusable for fungal enzymes because of hyperglycosylation.

Example 1. Production of Active Recombinant CBH I (rCBH I) in *Aspergillus awamori*

Construction of Modified CBH I Coding Sequence.

The coding sequence for *T. reesei* CBH I (SEQ. ID. NO: 4) was successfully inserted and expressed in *Aspergillus awamori* using the fungal expression vector pPFE2 (and pPFE1). Vectors pPFE1 and pPFE2 are *E. coli-Aspergillus* shuttle vectors, and contain elements required for maintenance in both hosts. They encode ampicillin resistance for selection in *E. coli* and

Zeocin resistance for selection in *Aspergillus*. The foregoing provided for the site-directed mutagenesis in *E. coli*, followed by expression of the new mutant proteins in *A. awamori*. The CBH I gene is under the control of the *A. awamori* glucoamylase promoter and includes the glucoamylase secretion signal peptide. In order to have the signal peptide properly cleaved during secretion, the construction of this plasmid required the addition, by PCR, of a NotI site and XbaI site on the coding sequence of CBH 1. The NotI site addition resulted in a change of the most N-terminal amino acid on the protein from glutamine to glycine. This glycine was subsequently changed back to the native glutamine in the pPFE2/CBHI construct, using site-directed mutagenesis PCR. This new construct was used to transform *A. awamori* and to express rCBH I, as confirmed by western blot analysis of culture supernatant. The rCBH I expressed in *A. awamori* tends to be over glycosylated as evidenced by the higher molecular weight observed on western blot analysis. Over-glycosylation of CBH I by *A. awamori* was confirmed by digestion of the recombinant protein with endoglycosidases. Following endoglycosidase H and F digestion, the higher molecular weight form of the protein collapses to a molecular weight similar to native CBH 1.

The vector pPFE2/CBHI requires a relatively long PCR reaction (8.2 kb) to make site-specific changes using the Stratagene Quik Change protocol. The PCR reaction was optimized as follows using a GeneAmp PCR System 2400, Perkin Elmer Corporation. The reaction mixture contained 50 ng of template DNA, 125 ng each of the sense and antisense mutagenic primers, 5 µl of Stratagene 10x cloned *Pfu* buffer, 200µM of each: dNTP, 5 mM $MgCl_2$ (total final concentration Of $MgCl_2$ is 7 mM); and 2.5 U *Pfu* Turbo DNA polymerase. The PCR reaction was carried out for 30 cycles, each consisting of one minute denaturation at 96C°, 1 minute annealing at 69°C, and 20-minute extension at 75°C. There is an initial denaturation for 2 minutes at 96°C and a final extension for 10 minutes at 75°C, followed by a hold at 4°C. Agarose gel electrophoresis, ethidium bromide staining, and visualization under UV transillumination were used to confirm the presence of a PCR product.

PCR products were digested with restriction enzyme DpnI, to degrade un-mutagenized parental DNA, and transformed into *E. coli* (Stratagene Epicurian Coli Supercompetent XL-1 Cells). $Amp^R$ colonies were picked from LB-$Amp^{100}$ plates and mutations were confirmed by DNA sequencing. Depending on scale, plasmid DNA was purified using the Qiagen QiaPrep Spin Miniprep Kit or the Promega Wizard Plus MaxiPrep DNA Purification System.

Transformation of *Aspergillus awamori* with *Trichoderma reesei* CBH I coding sequence.

*Aspergillus awamori* spore stocks were stored at -70°C in 20% glycerol 10% lactose. After thawing, 200 µL of spores were inoculated into 50 mL CM broth in each of eight baffled 250 mL Erlenmeyer flask. The cultures were grown at 28°C, 225 rpm for 48 h. The mycelia were removed by filtration with sterile Miracloth, Calbiochem, San Diego, CA, and washed thoroughly with sterile KCM. Approximately 10 g of washed mycelia were transferred to 50 mL KCM + 250 mg Novozym234 in a 250 mL baffled Erlenmeyer flask. The digestion mixture was incubated at 30°C, 80 rpm for 16-18 h. Spheroplasts were filtered through Miracloth into 50 mL conical centrifuge tubes, pelleted at 2000xg for 15 min and re-suspended in 0.7M KCl by gentle tituration with a 25 mL pipette. This procedure was repeated once. After a third pelleting, the spheroplasts were resuspended in 10 mL KC, pelleted and resuspended in 0.5 mL KC using a wide-bore pipet tip. The washed spheroplasts were transformed by adding 12.5 µL PCM and 5 µL DNA (≈0.5 µg/µL) to 50 µL of spheroplast in sterile 1.5 mL Eppendorf tubes. After incubation on ice for 45 minutes, 0.5 mL of room temperature PCM was added to the transformation mixture and was mixed by tituration with a wide bore pipet tip. The mixture was incubated at room temperature for 45 minutes. One milliliter of KC was added and mixed. The mixture was allocated between four tubes of molten CM top agar at 55°C, which were each poured over a 15 mL CM170 plate. The plates were incubated at 28°C for 2-3 days. Subsurface colonies were partially picked with a sterile wide bore pipet tip, exposing the remaining part of the colony to air and promoting rapid sporulation. After sporulation, spores were streaked onto several successive CM100 plates. After a monoculture was established, heavily sporulated plates were flooded with sterile spore suspension medium (20% glycerol, 10% lactose), the spores were suspended and aliquots were frozen at -70°C. Protein production was confirmed and followed by western blot using anti-CBH I monoclonal antibodies and the Novex Western Breeze anti-mouse chromogenic detection kit (Novex, San Diego, CA). Extracting genomic DNA using the YeaStar Genomic DNA Kit (Zymo Research, Orange, CA) and carrying out PCR with *pfu*-turbo DNA polymerase (Stratagene, La Jolla, CA) and *cbhI* primers confirmed insertion of the gene.

Production of Recombinant Enzyme.

For enzyme production, spores were inoculated into 50 mL CM maltose medium, pH 5.0, and grown at 32°C, 225 rpm in 250 mL baffled flasks. The cultures were transferred to 1.0 L of CM maltose in 2,800 mL Fernbach flasks and grown under similar conditions. For large-scale enzyme production (>1 mg), these cultures were transferred to 10-L CM maltose in a New Brunswick BioFlo3000 chemostat (10-L working volume) maintained at: 20% DO; pH 4.5; 25°C;

and 300 rpm. The culture was harvested by filtration through Miracloth after 2-3 days of growth. For the 10-L fermentation broth, the filtrate was concentrated and dia-filltered into 20mM sodium acetate pH 5.0 by tangential flow ultrafiltration with an Amicon DC30 concentrator equipped with a single 10,000 MWCO hollow fiber cartridge (l.1mm I.D., 2.4 $m^2$ surface area). The retentate from the 10-L concentration or the filtrate from smaller cultures was clarified in an Amicon DC-2 concentrator by tangential flow filtration with two 0.1 μm hollow fiber cartridges (1.1 mm I.D., 0.03 $m^2$ surface area, Millipore, Bedford, MA). The permeate was further concentrated with an Amicon CH-2 concentrator equipped with three 10,000 MWCO hollow fiber cartridges (1.1 mm I.D., 0.03 $m^2$ surface area). The final concentrate was sterile filtered through a 0.45 μm filter and stored at 4°C until used.

The recombinant CBH I protein, SEQ. ID NO.: 4, was purified by passing the concentrated culture broth over two or three CBinD900 cartridge columns (Novagen, Madison, WI) connected in series using a Pharmacia FPLC loading at 1.0 mL/min. (Amersham Pharmacia Biotech, Inc., Piscataway, NJ). The cartridges were equilibrated in 20 mM Bis-Tris pH 6.5 prior to loading and washed with the same buffer after loading. The bound rCBH I was eluted with 100% ethylene glycol (3 mL/column) by hand, using a syringe. The eluted rCBH I was concentrated in an Amicon 10 mL stirred cell using a 25 mm PM10 membrane to <2.0 mL and loaded onto a Pharmacia SuperDex200 16/60 size-exclusion column. The mobile phase was: 20 mM sodium acetate; 100 mM sodium chloride; and 0.02% sodium azide, pH 5.0 running at 1.0 mL/min. The eluted protein was concentrated by stirred cell and stored at 4°C. Concentration was determined by $A_{280}$ using the extinction coefficient and molecular weight calculated for individual proteins by the ProtParam tool on the ExPASy website (http://expasy.ch/tools/protparam.html. Below are the formulations for the various media described herein:

Clutterbuck's Salts (20X)

| | |
|---|---|
| $Na_2NO_3$ | 120.0 g/L |
| KCl | 10.4 g/L |
| $MgSO_4*7H_20$ | 10.4 g/L |
| $KH_2PO_4$ | 30.4 g/L |

CM- Yeast Extract- 5g/L

Tryptone- 5g/L

Glucose- 10g/L

Clutterbuck's Salts-50mL

Add above to 900mL $dH_20$, pH to 7.5, bring to 1000mL

CM Agar- CM+ 20g/L Agar

CMK CM Agar+ 0.7M KCl

CM100- CM + 100 g/mL Zeocin (Invitrogen, Carlsbad, CA)

CM 1070- CM+ 170 g/mL Zeocin

KCl- 0.7M KCl

KC- 0.7M KCl + 50mM CaCl2

KCM- 0.7M KCl + 10mM MOPS, pH 5.8

PCM 40% PEG 8000, 50mM $CaCl_2$, 10mM MOPS pH 5.8

Example 2. Production of Reduced Glycosylation rCBH 1: Sites N270A; N45A; and N384A.

rCBHl/pPFE2 has been optimized using site-directed mutagenesis to achieve expression of native molecular weight CBH I in *A. awamori*. The QuickChange SDM kit (Stratagene, San Diego, Ca) was used to make point mutations, switch amino acids, and delete or insert amino acids in the native CBH1 gene sequence. The Quick Change SDM technique was performed using thermotolerant *Pfu* DNA polymerase, which replicates both plasmid strands with high fidelity and without displacing the mutant oligonucleotide primers. The procedure used the polymerase chain reaction (PCR) to modify the cloned CBH1 DNA. The basic procedure used a supercoiled double stranded DNA (dsDNA) vector, with an insert of interest, and two synthetic oligonucleotide primers containing a desired mutation. The oligonucleotide primers, each complimentary to opposite strands of the vector, extend during temperature cycling by means of the polymerase. On incorporation of the primers, a mutated plasmid containing staggered nicks was generated. Following temperature cycling, the product was treated with a Dpn1 restriction enzyme. Dpn1 is specific for methylated and hemi-methylated DNA and thus digests the unmutated parental DNA template, selecting for the mutation-containing, newly-synthesized DNA. The nicked vector DNA, containing the desired mutations, was then transformed into *E. coli*. The small amount of template DNA required to perform this reaction, and the high fidelity of the *Pfu* DNA polymerase contribute to the high mutation efficiency and minimizes the potential for the introduction of random mutations. Three glycosylation-site amino acids on the protein surface were targeted for substitution of an alanine (A) residue in place of asparagine (N). Single site substitutions were successfully completed in the CBH I coding sequence at sites N45, N270, and N384, of SEQ. ID NO.: 4 by site-directed mutagenesis, and confirmed by DNA sequencing.

**Table 1.**

| CONSTRUCT | HOST | MW | $K_M$ | $V_{MAX}$ |
|---|---|---|---|---|
| *T. reesei* | none | 57.8 kDa | 1.94 | 0.746 |
| rCBH I wt cDNA# | *A. awamori* | 63.3 kDa | 2.14 | 0.668 |
| rCBH I wt genomic | *A. awamori* | 63.3 kDa | -- | -- |
| rCBH I N270A | *A. awamori* | 61.7 kDa | 2.25 | 0.489 |

As shown in **Table 1**, Western blot analysis of the supernatant, obtained from a single glycosylation-site mutant CBHIN270A (SEQ. ID NO.: 2) culture expressed in *A. awamori*, demonstrated that a decrease, to a lower molecular weight (61.7 kDa), in the amount of protein had occurred, as compared to the that in the wild type cDNA (63.3 kDa), and the wild type genomic DNA (63.3 kDa). These results demonstrate a reduction in the level of glycosylation in the reduced glycosylation mutant CBHIN270A, via expression in *A. awamori*. It is also shown, in the Table, that the CBHIN270A enzyme nearly retained its native enzymatic activity when assayed using the *p*NPL substrate. While not shown in the Table, variants CBHIN45A (SEQ. ID NO.: 1), and CBHI384A (SEQ. ID NO.: 3) have also demonstrated a reduction in amount of glycosylation and native activity when expressed from the heterologous host *A. awamori*.

Example 3. Production of Reduced Glycosylation rCBH 1: Double and Triple Mutants.

Double and triple combinations of this substitution have also been completed in the CBH I coding sequence (SEQ. ID NO.: 4) at sites N45, N270, and N384 by site-directed mutagenesis and confirmed by DNA sequencing. These double and triple-site constructs will also yield rCBH I enzymes with reduced glycosylation and, presumably, native activity.

Mutagenic Primers Used in Site-directed Mutagenesis PCR

Not1, Xbal insertion for vector construction

Mutagenic primers

C-terminal strand (Xbal): AGAGAGTCTAGACACGGAGCTTACAGGC

N-terminal strand (Notl):

AAAGAAGCGCGGCCGCGCCTGCACTCTCCAATCGG

Repair of NotI site to native sequence

Mutagenic primers

sense strand:GGCAAATGTGATTTCCAAGCGCCAGTCGGCCTGCACTCTCC

antisense strand:GGAGAGTGCAGGCCGACTGGCGCTTGGAAATCACATTTGCC

N45A glycosylation site mutation

Mutagenic primers

sense strand- GGACTCACGCTACGGCCAGCAGCACGAACTGC

antisense strand: GCAGTTCGTGCTGCTGGCCGTAGCGTGAGTCC

N270A glycosylation site mutation

Mutagenic primers

sense strand: CCCATACCGCCTGGGCGCCACCAGCTTCTACGGCCC

antisense strand: GGGCCGTAGAAGCTGGTGGCGCCCAGGCGGTATGGG

N384A glycosylation site mutation

Mutagenic primers

sense strand: GGACTCCACCTACCCGACAGCCGAGACCTCCTCCACACCCG

antisense strand:

CGGGTGTGGAGGAGGTCTCGGCTGTCGGGTAGGTGGAGTCC

The foregoing description is considered as illustrative only of the principles of the invention. Furthermore, since numerous modifications and changes will readily occur to those skilled in the art, it is not desired to limit the invention to the exact construction and process shown as described above. Accordingly, all suitable modifications and equivalents may be resorted to falling within the scope of the invention as defined by the claims which follow.

Claims

1. A method for making an active exoglucanase in a eukaryotic heterologous host, the method comprising reducing glycosylation of the exoglucanase, wherein reducing comprises replacing an N-glycosylation site amino acid residue with non-glycosyl accepting amino acid residue.

2. The method of claim 1, wherein the N-glycosylation site amino acid residues include asparagine 45, 270, or 384 of SEQ ID NO: 4 and the non-glycosyl accepting amino acid residue includes alanine.

3. The method of claims 1 wherein replacing comprises site-directed-mutagenesis.

4. The methods of claims 1 wherein the exoglucanase comprises a cellobiohydrolase.

5. An exoglucanase, comprising SEQ. ID. NO: 1.

6. An exoglucanase, comprising SEQ. ID. NO: 2.

7. An exoglucanase, comprising SEQ. ID. NO: 3.

8. An exoglucanase, comprising a combination of claims 5,6, or 7.

SEQUENCE LISTING

<110> Adney, William S.
Decker, Stephen R.
Lantz-McCarter, Suzanne
Baker, John O.
Vinzant, Todd B.
Nieves, Rafael A.
Himmel, Michael E.

<120> CELLOBIOHYDROLASE REDUCED GLYCOSYLATION VARIANTS: CBHIN45A; CBHIN270A; AND CBHIN384A

<130> HIMMEL NREL IR# 99-45

<140>
<141>

<160> 4

<170> PatentIn Ver. 2.0

<210> 1
<211> 496
<212> PRT
<213> Trichoderma reesei

<400> 1
Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
1 5 10 15

Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
20 25 30

Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Ala Ser Ser Thr
35 40 45

Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
50 55 60

Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
65 70 75 80

Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
85 90 95

Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
100 105 110

Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
115 120 125

Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu
130 135 140

Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr
145 150 155 160

Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys
165 170 175

Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp
180 185 190

Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser
195 200 205

Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
210 215 220

Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly
225 230 235 240

Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys
245 250 255

Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr Ser
260 265 270

Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu
275 280 285

Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr
290 295 300

Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser
305 310 315 320

Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala
325 330 335

Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe
340 345 350

Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp
355 360 365

Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asn
370 375 380

Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser
385 390 395 400

Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys Val
405 410 415

Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro
420 425 430

Ser Gly Gly Asn Pro Pro Gly Gly Asn Arg Gly Thr Thr Thr Thr Arg
435 440 445

Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His
450 455 460

Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala
465 470 475 480

Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
485 490 495

<210> 2
<211> 496
<212> PRT
<213> Trichoderma reesei

<400> 2
Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
1 5 10 15

Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
20 25 30

Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr
35 40 45

Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
50 55 60

Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
65 70 75 80

Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
85 90 95

Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
00 105 110

Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
115 120 125

Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu

130 135 140

Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr
145 150 155 160

Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys
165 170 175

Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp
180 185 190

Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser
195 200 205

Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
210 215 220

Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly
225 230 235 240

Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys
245 250 255

Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Ala Thr Ser
260 265 270

Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu
275 280 285

Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr
290 295 300

Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser
305 310 315 320

Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala
325 330 335

Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe
340 345 350

Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp
355 360 365

Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asn
370 375 380

Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser
385 390 395 400

Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys Val

405 410 415

Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro
420 425 430

Ser Gly Gly Asn Pro Pro Gly Gly Asn Arg Gly Thr Thr Thr Thr Arg
435 440 445

Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His
450 455 460

Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala
465 470 475 480

Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
485 490 495

<210> 3
<211> 496
<212> PRT
<213> Trichoderma reesei

<400> 3
Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
1 5 10 15

Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
20 25 30

Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr
35 40 45

Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
50 55 60

Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
65 70 75 80

Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
85 90 95

Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
00 105 110

Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
115 120 125

Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu
130 135 140

Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr
145 150 155 160

Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys
165 170 175

Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp
180 185 190

Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser
195 200 205

Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
210 215 220

Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly
225 230 235 240

Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys
245 250 255

Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr Ser
260 265 270

Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu
275 280 285

Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr
290 295 300

Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser
305 310 315 320

Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala
325 330 335

Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe
340 345 350

Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp
355 360 365

Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Ala
370 375 380

Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser
385 390 395 400

Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys Val
405 410 415

Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro
420 425 430

Ser Gly Gly Asn Pro Pro Gly Gly Asn Arg Gly Thr Thr Thr Thr Arg
435 440 445

Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His
450 455 460

Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala
465 470 475 480

Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
485 490 495

<210> 4
<211> 496
<212> PRT
<213> Trichoderma reesei

<400> 4
Gln Ser Ala Cys Thr Leu Gln Ser Glu Thr His Pro Pro Leu Thr Trp
1 5 10 15

Gln Lys Cys Ser Ser Gly Gly Thr Cys Thr Gln Gln Thr Gly Ser Val
20 25 30

Val Ile Asp Ala Asn Trp Arg Trp Thr His Ala Thr Asn Ser Ser Thr
35 40 45

Asn Cys Tyr Asp Gly Asn Thr Trp Ser Ser Thr Leu Cys Pro Asp Asn
50 55 60

Glu Thr Cys Ala Lys Asn Cys Cys Leu Asp Gly Ala Ala Tyr Ala Ser
65 70 75 80

Thr Tyr Gly Val Thr Thr Ser Gly Asn Ser Leu Ser Ile Gly Phe Val
85 90 95

Thr Gln Ser Ala Gln Lys Asn Val Gly Ala Arg Leu Tyr Leu Met Ala
00 105 110

Ser Asp Thr Thr Tyr Gln Glu Phe Thr Leu Leu Gly Asn Glu Phe Ser
115 120 125

Phe Asp Val Asp Val Ser Gln Leu Pro Cys Gly Leu Asn Gly Ala Leu
130 135 140

Tyr Phe Val Ser Met Asp Ala Asp Gly Gly Val Ser Lys Tyr Pro Thr
145 150 155 160

Asn Thr Ala Gly Ala Lys Tyr Gly Thr Gly Tyr Cys Asp Ser Gln Cys
165 170 175

Pro Arg Asp Leu Lys Phe Ile Asn Gly Gln Ala Asn Val Glu Gly Trp
180 185 190

Glu Pro Ser Ser Asn Asn Ala Asn Thr Gly Ile Gly Gly His Gly Ser
195 200 205

Cys Cys Ser Glu Met Asp Ile Trp Glu Ala Asn Ser Ile Ser Glu Ala
210 215 220

Leu Thr Pro His Pro Cys Thr Thr Val Gly Gln Glu Ile Cys Glu Gly
225 230 235 240

Asp Gly Cys Gly Gly Thr Tyr Ser Asp Asn Arg Tyr Gly Gly Thr Cys
245 250 255

Asp Pro Asp Gly Cys Asp Trp Asn Pro Tyr Arg Leu Gly Asn Thr Ser
260 265 270

Phe Tyr Gly Pro Gly Ser Ser Phe Thr Leu Asp Thr Thr Lys Lys Leu
275 280 285

Thr Val Val Thr Gln Phe Glu Thr Ser Gly Ala Ile Asn Arg Tyr Tyr
290 295 300

Val Gln Asn Gly Val Thr Phe Gln Gln Pro Asn Ala Glu Leu Gly Ser
305 310 315 320

Tyr Ser Gly Asn Glu Leu Asn Asp Asp Tyr Cys Thr Ala Glu Glu Ala
325 330 335

Glu Phe Gly Gly Ser Ser Phe Ser Asp Lys Gly Gly Leu Thr Gln Phe
340 345 350

Lys Lys Ala Thr Ser Gly Gly Met Val Leu Val Met Ser Leu Trp Asp
355 360 365

Asp Tyr Tyr Ala Asn Met Leu Trp Leu Asp Ser Thr Tyr Pro Thr Asn
370 375 380

Glu Thr Ser Ser Thr Pro Gly Ala Val Arg Gly Ser Cys Ser Thr Ser
385 390 395 400

Ser Gly Val Pro Ala Gln Val Glu Ser Gln Ser Pro Asn Ala Lys Val
405 410 415

Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Gly Ser Thr Gly Asn Pro
420 425 430

Ser Gly Gly Asn Pro Pro Gly Gly Asn Arg Gly Thr Thr Thr Thr Arg
435 440 445

Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly Pro Thr Gln Ser His
450 455 460

Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly Pro Thr Val Cys Ala
465 470 475 480

Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr Tyr Ser Gln Cys Leu
485 490 495